Revised June 19, 2014

Catalogue no. 99-011-X2011002
ISBN: 978-1-100-23690-2

Technical Report

Aboriginal Peoples Technical Report

National Household Survey, 2011

Statistics Canada Statistique Canada

Canada

How to obtain more information

For information about this product or the wide range of services and data available from Statistics Canada, visit our website, www.statcan.gc.ca.

You can also contact us by

email at infostats@statcan.gc.ca

telephone, from Monday to Friday, 8:30 a.m. to 4:30 p.m., at the following numbers:

- Statistical Information Service 1-800-263-1136
- National telecommunications device for the hearing impaired 1-800-363-7629
- Fax line 1-877-287-4369

Depository Services Program

- Inquiries line 1-800-635-7943
- Fax line 1-800-565-7757

To access this product

This product, Catalogue no. 99-011-X2011002, is available free in electronic format. To obtain a single issue, visit our website, www.statcan.gc.ca, and browse by "Key resource" > "Publications."

Standards of service to the public

Statistics Canada is committed to serving its clients in a prompt, reliable and courteous manner. To this end, this agency has developed standards of service that its employees observe. To obtain a copy of these service standards, please contact Statistics Canada toll-free at 1-800-263-1136. The service standards are also published at www.statcan.gc.ca under "About us" > "The agency" > "Providing services to Canadians."

Published by authority of the Minister responsible for Statistics Canada

© Minister of Industry, 2014

All rights reserved. Use of this publication is governed by the Statistics Canada Open Licence Agreement (www.statcan.gc.ca/reference/licence-eng.html).

Cette publication est aussi disponible en français.

Note of appreciation

Canada owes the success of its statistical system to a long-standing partnership between Statistics Canada, the citizens of Canada, its businesses, governments and other institutions. Accurate and timely statistical information could not be produced without their continued co-operation and goodwill.

Standard symbols

The following symbols are used in Statistics Canada publications:

.	not available for any reference period
..	not available for a specific reference period
...	not applicable
0	true zero or a value rounded to zero
0^s	value rounded to 0 (zero) where there is a meaningful distinction between true zero and the value that was rounded
^{p}	preliminary
^{r}	revised
x	suppressed to meet the confidentiality requirements of the *Statistics Act*
^{E}	use with caution
F	too unreliable to be published
*	significantly different from reference category ($p < 0.05$)

Table of contents

Page

Introduction **5**

1 Concepts and variables **6**
1.1 Aboriginal concepts 6
1.1.1 Aboriginal group (Question 18) 6
1.1.2 The 2011 NHS Registered or Treaty Indian status question 7
1.1.3 The 2011 NHS First Nation/Indian band membership question 8
1.1.4 Aboriginal ancestry (Question 17) 9
1.2 Aboriginal variables 11
1.2.1 Aboriginal identity derived 11
1.2.2 Other possible derived variables 12
1.3 Aboriginal geography variables 13
1.3.1 Aboriginal area of residence – On reserve 13
1.3.2 Inuit area of residence 13

2 Data collection **14**
2.1 General collection methods 14
2.1.1 Early enumeration and collection on Indian reserves, Métis settlements and Inuit communities 15
2.2 Census and NHS response rates (collection) 15
2.3 Questionnaires and instruction guides 16
2.3.1 Questionnaire languages 16

3 Data processing **17**
3.1 Receipt, registration and storage 17
3.2 Imaging and data capture from paper questionnaires 17
3.3 Edits and failed edit follow-up 18
3.4 Coding 18
3.4.1 Coding of the First Nation/Indian band membership write-in question 19
3.5 Edit and imputation 19
3.6 Edit and imputation of Aboriginal variables 19
3.7 Weighting 20

4 Data dissemination **22**
4.1 Analytical products 22
4.1.1 Analytical documents 22
4.2 Data products 22
4.2.1 NHS Profiles 22
4.2.2 NHS Data Tables – Aboriginal Peoples 23
4.2.3 NHS Focus on Geography Series 23
4.3 Reference products 23
4.3.1 NHS Dictionary 23
4.3.2 Aboriginal Peoples Reference Guide 24
4.3.3 Incompletely enumerated Indian reserves and Indian settlements 24
4.3.4 Custom products 24

5 Data quality assessment and indicators 25
5.1 Sources of error 25
5.1.1 Sampling error 25
5.1.2 Non-sampling error 26
5.2 Data suppression related to confidentiality and data quality 27
5.2.1 Data suppression related to confidentiality (non-disclosure) 27
5.2.2 Data quality indicators 29
5.2.3 Other occurrences when data are suppressed or not available 29
5.2.4 Data availability from the NHS for census subdivisions 30
5.3 Coverage 34
5.3.1 Net undercoverage error for participating reserves 34
5.3.2 Coverage error for incompletely enumerated reserves and settlements 36

6 Comparability 38
6.1 Historical comparability 38
6.1.1 Differences in wording and format of Aboriginal questions 38
6.1.2 Legislative changes 38
6.1.3 Differences in the list set of incompletely enumerated reserves 39
6.1.4 Changes made to the census subdivision types associated with 'on reserve' population 39
6.1.5 Processing and methodological changes 40
6.2 Undercoverage of population in participating Indian reserves and settlements 41
6.3 Comparability with other sources 41
6.3.1 Comparability with the Indian Register 41
6.3.2 Comparability with Demosim projections 42

7 Conclusions 44

Introduction

The 2011 National Household Survey (NHS) is a new survey which replaces the long census questionnaire, more widely known as the Census Form 2B.

The NHS covers all persons who usually live in Canada, in the provinces and the territories. It includes persons who live on Indian reserves and settlements, permanent residents, non-permanent residents such as refugee claimants, holders of work or study permits, and members of their families living with them.[1]

The 2011 *Aboriginal Peoples Technical Report, National Household Survey* aims to inform users about the conceptual framework and definitions used to gather the 2011 NHS data on Aboriginal people, how the data are collected, the products used to disseminate the information and the data quality aspects relevant to this topic.

Section 1 describes the various concepts and variables that can be used to define the Aboriginal population. Section 2 and Section 3 describe the data collection and data processing procedures. Section 4 presents the various NHS analytical, data and reference products related to Aboriginal people. Confidentiality and data quality are discussed in Section 5. Section 6 discusses data comparability over time and with other data sources. Section 7 presents some final remarks about data comparability and data quality of the NHS data for Aboriginal people.

1. For more information on the NHS target population, refer to the *National Household Survey User Guide*, Catalogue no. 99-001-XWE2011001 (www12.statcan.gc.ca/nhs-enm/2011/ref/nhs-enm_guide/guide_1-eng.cfm).

1 Concepts and variables

There are various ways to define the Aboriginal population based on the four related questions asked in the National Household Survey (NHS) depending on the focus and the requirements of the data user.

The definitions of NHS terminology, variables and concepts are presented here and appear in the *National Household Survey Dictionary*, Catalogue 99-000-X. Users should refer to the *National Household Survey Dictionary* for full definitions and additional remarks related to any concepts, such as information on direct and derived variables and their respective universe.

1.1 Aboriginal concepts

The four questions used on the NHS to identify Aboriginal peoples yield different concepts for defining different Aboriginal populations. Four commonly used concepts include:

- Aboriginal ancestry
- Aboriginal identity
- First Nation/Indian band membership
- Registered or Treaty Indian status.

The population estimates, at the Canada level, from the 2011 National Household Survey using the different definitions are shown below.

Aboriginal ancestry[2]	1,836,035
Aboriginal identity[3]	1,400,685
First Nation/Indian band membership	675,485
Registered or Treaty Indian status	697,505

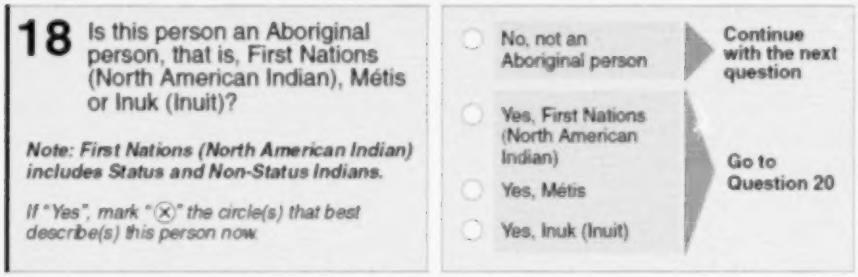
1.1.1 Aboriginal group (Question 18)

Aboriginal group refers to whether a person belongs to the Aboriginal peoples of Canada. In previous censuses, respondents were asked to report if they were an Aboriginal person, that is, First Nations (North American Indian), Métis, or Inuk (Inuit). The 1996 Census was the first census to ask the Aboriginal group question, using the wording: 'Is this person an Aboriginal person, that is, North American Indian, Métis or Inuit (Eskimo)?' This same wording was used in this question on the 2001 and 2006 censuses.

For the 2011 NHS questionnaires (N1 and N2), changes were introduced to the terminology in this question. 'North American Indian' was replaced by 'First Nations (North American Indian)' and 'Inuit (Eskimo)' was replaced by 'Inuk (Inuit)' in both the question and the response categories. An instruction was added '**Note: First Nations (North American Indian) includes Status and Non-Status Indians.**' See Figure 1 for the wording of the 2011 NHS Aboriginal group question.

2. Included are single and multiple ancestry responses. A single ancestry response is when a respondent provides one Aboriginal ancestry only. A multiple response is given when a respondent provides two or more ancestries with at least one being an Aboriginal ancestry.
3. Included are persons who reported identifying with at least one Aboriginal group, and/or those who reported being a Registered or Treaty Indian, as defined by the *Indian Act* of Canada, and/or those who reported they were members of a First Nation or Indian band.

Figure 1 The 2011 NHS Aboriginal group question (N1 and N2 questionnaires)


18 Is this person an Aboriginal person, that is, First Nations (North American Indian), Métis or Inuk (Inuit)?
Note: First Nations (North American Indian) includes Status and Non-Status Indians.
If "Yes", mark "⊗" the circle(s) that best describe(s) this person now.
No, not an Aboriginal person
Continue with the next question
Yes, First Nations (North American Indian)
Yes, Métis
Yes, Inuk (Inuit)
Go to Question 20

Source: Statistics Canada, National Household Survey, 2011.

The following information was provided to respondents in the *National Household Survey Guide*.

- Answer this question regardless of whether or not this person is an Aboriginal person of North America.
- Aboriginal people are usually those with ancestors who resided in North America prior to European contact and who identify with one of the three Aboriginal groups listed on the questionnaire—First Nations (North American Indian), Métis and Inuit.
- Persons who consider themselves to be East Indian or Asian Indian, or who have ethnic roots on the subcontinent of India, should respond 'No' to this question.
- Individuals who refer to themselves as Métis in the context of mixed ancestry, but who do not have North American Aboriginal ancestry—for example, those from Africa, the Caribbean and South America—should respond 'No.'

1.1.2 The 2011 NHS Registered or Treaty Indian status question

Registered or Treaty Indian status refers to whether or not a person reported being a Registered or Treaty Indian. Registered Indian refers to persons who are registered under the *Indian Act* of Canada. Treaty Indians are persons who belong to a First Nation/Indian band that signed a treaty with the Crown. Registered or Treaty Indians are sometimes also called Status Indians.[4]

Data on Registered or Treaty Indian status were first collected in the 1981 Census using response categories in the ethnic origin question. Within the ethnic origin question, mark-in options under a sub-heading for 'Native Peoples' included 'Status or registered Indian' and 'Non-status Indian.'

In the 1986 Census, data were not collected specifically for Registered or Treaty Indian status. That year, the ethnic origin question included a mark-in option for 'North American Indian' that was meant to include both status and non-status Indians together.

In 1991, a new census question on Registered Indians had two components. In the first part of the question, respondents were asked about their registration status, while the second part of the question dealt with band membership. The first part of the 1991 Census registration status question asked 'Is this person a **registered Indian** as defined by the *Indian Act* of Canada?'

4. For more information, refer to www.statcan.gc.ca/concepts/definitions/aboriginal-autochtone4-eng.htm.

A change from 1991, the 1996 Census Registered or Treaty Indian status question asked only about registration or treaty status, while band membership was dealt with in a separate question. The 1996 Census question wording, 'Is this person a Treaty Indian or a Registered Indian as defined by the *Indian Act* of Canada?' differed slightly from the wording used in previous censuses. Prior to 1996, the term 'treaty' was not included in the question. It was added in 1996 at the request of individuals from the Western provinces, where the term was more widely used. This same question was used in the 2001 and 2006 censuses.

In the 2011 NHS question, the term 'Status Indian' was included to clarify the question, since this term is used more extensively than 'Registered or Treaty Indian' across Canada. See Figure 2 for the wording of the 2011 NHS Registered or Treaty Indian status question.

Figure 2 The 2011 NHS Registered or Treaty Indian status question

20	Is this person a Status Indian (Registered or Treaty Indian as defined by the *Indian Act* of Canada)?	○ No ○ Yes, Status Indian (Registered or Treaty)

Source: Statistics Canada, National Household Survey, 2011.

The following information was provided to respondents in the *National Household Survey Guide*.

Select the circle 'Yes, Status Indian (Registered or Treaty)' for persons who:

- are Registered Indians under the *Indian Act*
- are Treaty Indians, only if they are Registered Indians under the *Indian Act*
- have become registered since June 1985, when Bill C-31 changed the *Indian Act*.

All other persons should mark 'No' including persons who may be entitled to register under provisions of the *Indian Act*, but for some reason have not.

1.1.3 The 2011 NHS First Nation/Indian band membership question

Membership in a First Nation/Indian band refers to whether a person is a member of a First Nation or Indian band of Canada.

Aboriginal Affairs and Northern Development Canada (AANDC) defines a band as a body of Indians for whose collective use and benefit lands have been set apart or money is held by the Crown, or declared to be a band for the purpose of the *Indian Act*.

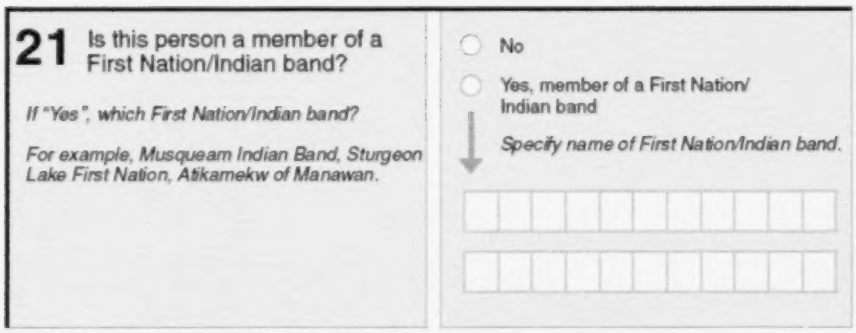
Data on Indian band membership were first collected on the 1971 Census. In 1971, data were collected using response categories in the ethnic origin question. Within the ethnic origin question, mark-in options included 'Native Indian – Band' and 'Native Indian – Non-band.' Data were not available from the census on Indian band membership again until 1991.

In 1991, a new census question on membership in a First Nation or Indian band had two components. In the first part of the question, respondents were asked about their registration status, while the second part of the question dealt with band membership.

In 1996, a direct question was developed for the census and asked 'Is this person a member of an Indian Band/First Nation?' This question was placed before the Registered or Treaty Indian status question. The same membership question and the same order of questions were used again in the 2001 and 2006 censuses.

In the 2011 NHS, the terms 'Indian band' and 'First Nation' were reversed from the 2006 Census version of the question, as a result of preferences expressed during testing of the question. The format of the question was comprised of two mark-in circles and a write-in box. There was a mark-in circle for the option 'No,' and a mark-in circle for the option 'Yes, member of a First Nation/Indian band,' followed by a write-in box. This write-in box contained two 11-segment lines for the option 'Specify name of First Nation/Indian band.' on the paper questionnaire, but writing above, between or below the lines was also captured when the questionnaire was scanned. In the Internet version of the question, there was space for 60 letters in the answer. The question also contains three examples of First Nations/Indian bands, one that contains 'Indian band' in the name, one that uses 'First Nation' in the name and one that uses neither. These examples were added because some participants during testing on reserve were answering with the name of their Nation/Tribe rather than the exact name of their First Nation/Indian band. See Figure 3 for the wording of the 2011 NHS First Nation/Indian band membership question.

Figure 3 The 2011 NHS First Nation/Indian band membership question


21 Is this person a member of a First Nation/Indian band?
If "Yes", which First Nation/Indian band?
For example, Musqueam Indian Band, Sturgeon Lake First Nation, Atikamekw of Manawan.
No
Yes, member of a First Nation/ Indian band
Specify name of First Nation/Indian band.

Source: Statistics Canada, National Household Survey, 2011.

The following information was provided to respondents in the *National Household Survey Guide*.

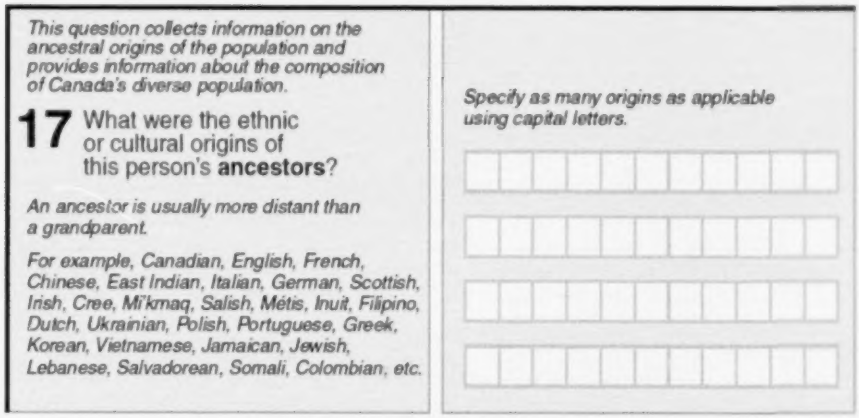
- A First Nation/Indian band is a group of people for whom lands have been set apart and/or money is held by the Crown.
- A member of a First Nation/Indian band is an individual who is recognized as being a member of a First Nation/Indian band, as defined by either the band itself or the *Indian Act*.
- Individuals should report their First Nation/Indian band affiliation rather than their tribal affiliation—for example, 'Chemawawin First Nation Band' instead of 'Cree.'

1.1.4 Aboriginal ancestry (Question 17)

Aboriginal ancestry refers to whether a person reported ancestry associated with the Aboriginal peoples of Canada in the ethnic origin question. Ethnic origin refers to the ethnic or cultural origins of the respondent's ancestors, an ancestor being usually more distant than a grandparent. A person can have more than one ethnic or cultural origin.

From 1981 to the present, ethnic origins, including Aboriginal ancestry, have been defined by descent from both the mother's and the father's side. The ethnic origin question has allowed for the reporting of single and multiple responses. See Figure 4 for the wording of the 2011 NHS Ethnic Origin question.

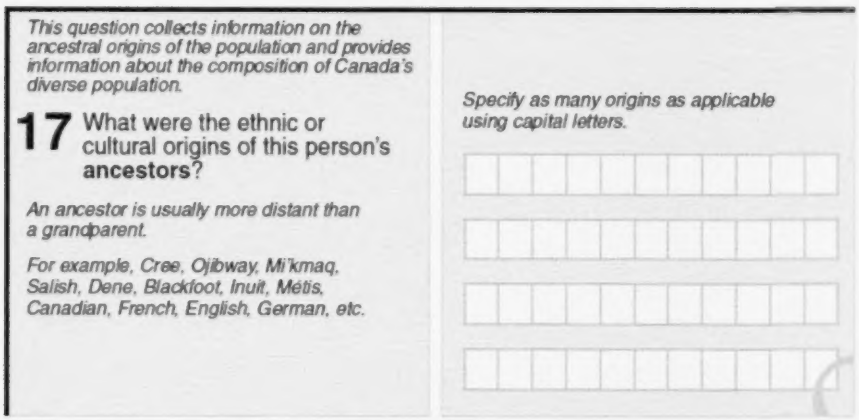
Figure 4 The 2011 NHS Ethnic Origin question (N1 questionnaire)


This question collects information on the ancestral origins of the population and provides information about the composition of Canada's diverse population.
17 What were the ethnic or cultural origins of this person's ancestors?
An ancestor is usually more distant than a grandparent.
For example, Canadian, English, French, Chinese, East Indian, Italian, German, Scottish, Irish, Cree, Mi'kmaq, Salish, Métis, Inuit, Filipino, Dutch, Ukrainian, Polish, Portuguese, Greek, Korean, Vietnamese, Jamaican, Jewish, Lebanese, Salvadorean, Somali, Colombian, etc.
Specify as many origins as applicable using capital letters.

Source: Statistics Canada, National Household Survey, 2011.

Note that while the ancestry (ethnic origin) question asked in the 2011 National Household N2 questionnaire (Figure 5) used for Aboriginal communities is the same as that asked in the N1 questionnaire for the general population, the list of examples are different. The N2 questionnaire is the Northern and Reserves questionnaire.

Figure 5 The 2011 NHS Ethnic Origin question (N2 questionnaire)


This question collects information on the ancestral origins of the population and provides information about the composition of Canada's diverse population.
17 What were the ethnic or cultural origins of this person's ancestors?
An ancestor is usually more distant than a grandparent.
For example, Cree, Ojibway, Mi'kmaq, Salish, Dene, Blackfoot, Inuit, Métis, Canadian, French, English, German, etc.
Specify as many origins as applicable using capital letters.

Source: Statistics Canada, National Household Survey, 2011.

The following information was provided to respondents in the *National Household Survey Guide*.

- This question refers to the ethnic or cultural origin or origins of a person's ancestors. Other than Aboriginal persons, most people can trace their origins to their ancestors who first came to this continent. Ancestry should not be confused with citizenship or nationality.
- For all persons, report the specific ethnic or cultural group or groups to which their ancestors belonged, not the language they spoke. For example, report 'Haitian' rather than 'French,' or 'Austrian' rather than 'German.'
- For persons of East Indian or South Asian origins, report a specific origin or origins. Do not report 'Indian.' For example, report 'East Indian from India,' 'East Indian from Guyana,' or indicate the specific group, such as 'Punjabi' or 'Tamil.'
- For persons with Aboriginal ancestors, report a specific origin or origins. For example, report 'Cree,' 'Mi'kmaq,' 'Ojibway,' 'Métis,' or 'North American Indian.' Do not report 'Indian.'

1.2 Aboriginal variables

1.2.1 Aboriginal identity derived

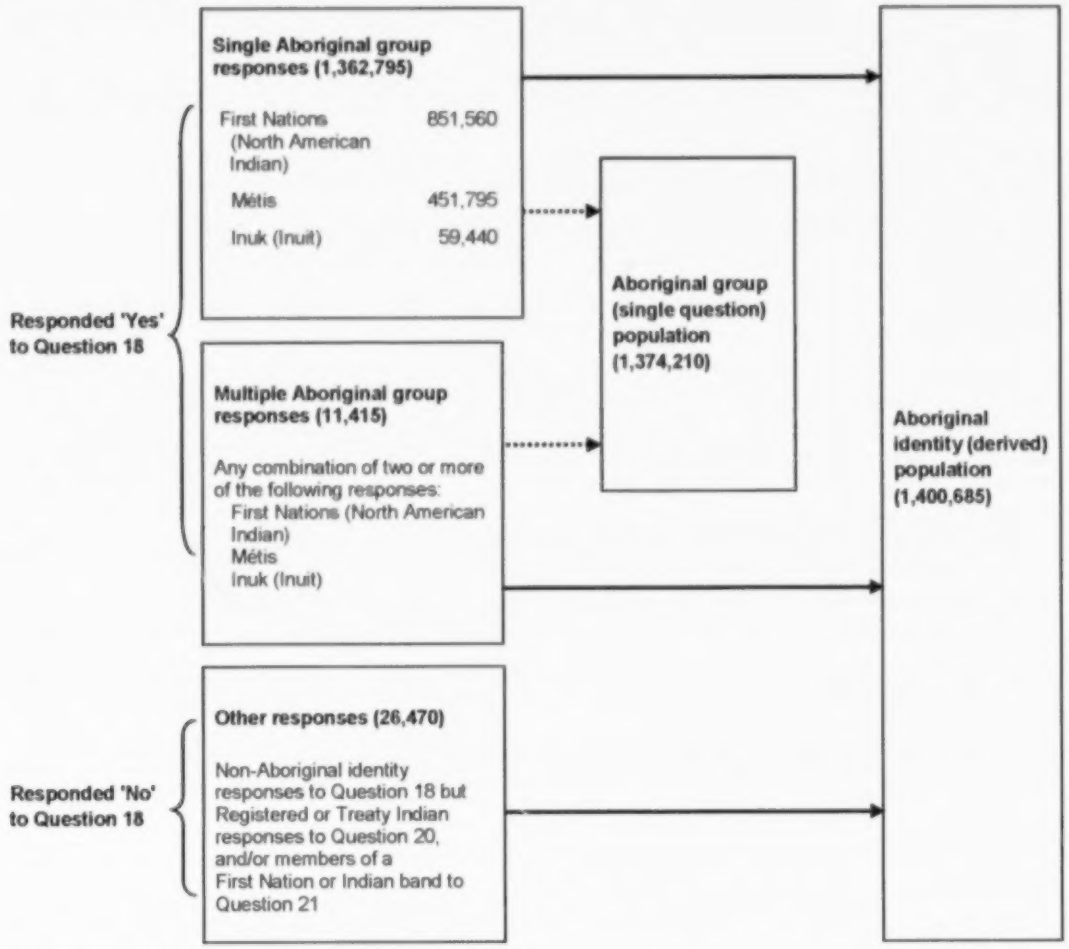
The derived Aboriginal identity concept refers to those persons who reported identifying with at least one Aboriginal group in Question 18, that is, First Nations (North American Indian), Métis or Inuk (Inuit), and/or those who reported being a Registered/Treaty Indian in Question 20, as defined by the *Indian Act* of Canada, and/or those who reported they were members of a First Nation/Indian band in Question 21.

The Aboriginal identity population is derived from three questions (Questions 18, 20 and 21).

Included in the derived Aboriginal identity population are the following (see also Figure 6):

Total Aboriginal identity population
- 'First Nations (North American Indian)' single response
- 'Métis' single response
- 'Inuk (Inuit)' single response
- Multiple Aboriginal responses
- Aboriginal responses not included elsewhere (includes those who reported themselves as being Registered or Treaty Indians and/or First Nations/Indian band members without reporting Aboriginal identity)

Figure 6 Aboriginal identity derived


Single Aboriginal group responses (1,362,795)
First Nations (North American Indian) 851,560
Métis 451,795
Inuk (Inuit) 59,440
Responded 'Yes' to Question 18
Multiple Aboriginal group responses (11,415)
Any combination of two or more of the following responses:
First Nations (North American Indian)
Métis
Inuk (Inuit)
Aboriginal group (single question) population (1,374,210)
Aboriginal identity (derived) population (1,400,685)
Other responses (26,470)
Non-Aboriginal identity responses to Question 18 but Registered or Treaty Indian responses to Question 20, and/or members of a First Nation or Indian band to Question 21
Responded 'No' to Question 18

Source: Statistics Canada, National Household Survey, 2011.

1.2.2 Other possible derived variables

It is possible to derive other definitions of the Aboriginal population, using different combinations of NHS questions. For example, the information from Question 18 on Aboriginal group and Question 20 on Registered/Treaty Indian status can be used to derive the First Nations (North American Indian) identity population with Registered or Treaty Indian status.

Other examples are:

- First Nations (North American Indian) identity population without Registered or Treaty Indian status (non-status)
- First Nations (North American Indian) identity population with Indian band membership
- Registered or Treaty Indian population with Indian band membership.

Other combinations of Aboriginal variables that can correspond to users' data needs and purposes are also possible.

1.3 Aboriginal geography variables

1.3.1 Aboriginal area of residence – On reserve

'Area of residence – On reserve' refers to the following geographic areas: Indian reserves and settlements (referred to as 'On reserve'), and all other areas (referred to as 'Off reserve'). These geographic areas can be used to show where people, primarily Registered or Treaty Indians and First Nations people, reside. The 2011 NHS total 'On reserve' population was 360,620, with about nine out of ten being First Nations people.

'On reserve' includes six census subdivision (CSD) types legally affiliated with First Nations or Indian bands, i.e., Indian reserve (IRI), Indian settlement (S-É) (except for the five Yukon settlements of Champagne Landing 10, Klukshu, Two and One-Half Mile Village, Two Mile Village and Kloo Lake), Indian government district (IGD), *terres réservées aux Cris* (TC), *terres réservées aux Naskapis* (TK) and Nisga'a land (NL), as well as the northern village of Sandy Bay in Saskatchewan.[5]

1.3.2 Inuit area of residence

'Area of residence: Inuit Nunangat' refers to the four regions of Inuit Nunangat and 'outside Inuit Nunangat.'

Inuit Nunangat is the homeland of Inuit of Canada. It includes the communities located in the four Inuit regions: Nunatsiavut (Northern coastal Labrador), Nunavik (Northern Quebec), the territory of Nunavut and the Inuvialuit region of the Northwest Territories. These regions collectively encompass the area traditionally occupied by Inuit in Canada.[6]

The 2011 NHS Inuit Nunangat population is 52,115 of which eight in ten are Inuit. In the four regions of Inuit Nunangat, the population breakdown is as follows:

Table 1 Inuit identity population by Inuit area of residence, 2011 NHS

Region	Total population	Percentage of population with Inuit identity (%)
Inuit Nunangat	52,115	83
Nunatsiavut	2,615	89
Nunavik	12,060	89
Nunavut	31,695	85
Inuvialuit region	5,745	58

Source: Statistics Canada, National Household Survey, 2011.

5. For more information on the 'On reserve' area of residence, refer to www12.statcan.gc.ca/nhs-enm/2011/ref/dict/pop150-eng.cfm.
6. For more information on 'Inuit Nunangat' area of residence, refer to www12.statcan.gc.ca/nhs-enm/2011/ref/dict/pop149-eng.cfm.

2 Data collection

The 2011 Census contained 10 questions and was conducted in May 2011.[7] The information previously collected by the long-form census questionnaire was collected as part of the new National Household Survey (NHS). The NHS questionnaire covered most of the same topics as the 2006 Census long form.

Between May and August 2011, Statistics Canada conducted the National Household Survey (NHS) for the first time. This voluntary, self-administered survey was introduced as a replacement for the Census Form 2B, more widely known as long census questionnaire. Roughly 4.5 million dwellings across Canada were selected for the NHS, representing about one-third of all private dwellings. The NHS covered all persons who usually live in Canada, in the provinces and the territories. It included persons who live on Indian reserves or settlements, permanent residents, non-permanent residents such as refugee claimants, holders of work or study permits, and members of their families living with them.

The survey excluded persons living in institutional collective dwellings such as hospitals, nursing homes and penitentiaries; Canadian citizens living in other countries; and full-time members of the Canadian Forces stationed outside Canada. Also excluded were persons living in non-institutional collective dwellings such as work camps, hotels and motels, and student residences. Therefore, Aboriginal people living outside Canada, in collective dwellings or institutions or who were homeless on Census Day were excluded from the survey.

For more information about the reference date and the census collection, refer to *About the census* [7]

2.1 General collection methods

Respondents living outside Indian reserves and settlements, Métis settlements and Inuit as well as remote communities had three response options:

- **An online questionnaire:** Occupants of dwellings selected for the NHS could answer the NHS questionnaire online either immediately after finishing the online census questionnaire, or later using the secure access code provided on the NHS paper questionnaire or reminder letter.

- **A paper questionnaire:** Occupants of most dwellings selected for the NHS who did not respond online to the census could complete a printed NHS questionnaire sent by mail or dropped off by a Statistics Canada enumerator in early June 2011.

- **An interview with a Statistics Canada enumerator:** This method was used in non-response follow-up. It was also offered to respondents who wanted to complete their questionnaire by telephone by calling the survey's help line.

The sampling fraction varied with the questionnaire delivery mode. For the mail delivery mode, about 3 dwellings in 10 (29%) received an NHS questionnaire. For the enumerator delivery mode, the sampling fraction was 1 in 3 dwellings (33%). However, in cases where it was necessary to reach households in remote areas or on Indian reserves, where only the interview response mode was offered, all households (100%) were invited to participate in the NHS.

7. For more information on the 2011 Census, refer to http://www12.statcan.gc.ca/census-recensement/2011/ref/about-apropos/index-eng.cfm

2.1.1 Early enumeration and collection on Indian reserves, Métis settlements and Inuit communities

In these areas, the 2011 Census and NHS enumerations were conducted at the same time. In some northern and remote areas of Canada, enumeration was conducted before Census Day, in February, March and April 2011, in order to reach respondents before they migrated to hunting and fishing camps for the summer, and because of seasonal climate variations, or other travel impediments. When enumeration took place before Census Day (May 10, 2011), the reference date used was the date on which the household was enumerated.

Overall for the NHS, approximately 5,000 crew leaders and assistants were hired to recruit, train and supervise around 30,000 census enumerators. Canvasser enumeration also took place on Indian reserves and settlements, and Métis settlements in southern areas of the country primarily from May 2 to July 29, 2011, but extended into early August in some areas. No non-response follow-up (NRFU) took place in areas of early enumeration and reserve enumeration, as completeness (for example, dwelling coverage) was verified during the enumeration process.

For enumeration on Indian reserves and settlements, certain procedures were followed to contact chief and council prior to collection and obtain permission to conduct the census and the National Household Survey. As much as possible, community members were hired to help with enumeration as crew leaders, enumerators, guides or translators.

2.2 Census and NHS response rates (collection)

The 2011 Census response rate was 97.1% at the Canada level but varied by province and territory. The highest rate was 97.7% in Prince Edward Island and the lowest rate was 91.4% for Nunavut. (Refer to www12.statcan.gc.ca/census-recensement/2011/ref/about-apropos/rates-taux-eng.cfm for the rates by provinces and territories.)

The NHS, being a voluntary survey, was not expected to achieve the same level of participation as the census. Unweighted response rates are calculated from final data, following the completion of data processing and data quality verification. The final status of a dwelling as respondent, non-respondent, or out of scope for the National Household Survey is done as part of data processing. The unweighted response rates are then calculated as the number of sampled private dwellings that returned a questionnaire divided by the number of sampled private dwellings classified as occupied by field staff. The unweighted response rate was 68.6% for the National Household Survey for Canada, all collection methods combined.

While this is similar to the response rate for other voluntary surveys conducted by Statistics Canada, it is lower than the 2011 Census response rate. The NHS response rates varied by province and territory; the highest unweighted rate was 83.9% in Northwest Territories and the lowest was 60.4% in Prince Edward Island.

Weighted response rates are based on final design weights of the NHS. Dwellings that responded to the NHS prior to a fixed date in collection have their initial design weight based on the sampling fraction in their area. After this fixed date, in order to limit the non-response bias as much as possible with the available resources, the NHS focused collection operations on a subsample of remaining non respondents. Within this subsample, the design weights were increased to reflect this change. The weighted response rates were then calculated as the weighted number of sampled private dwellings that returned a questionnaire divided by the weighted number of sampled private dwellings classified as occupied by field staff. The weighted response rate was 77.2% for the National Household Survey for Canada, all collection methods combined.

NHS national and provincial/territorial weighted and unweighted response rates, as well as those for census subdivisions (e.g., municipalities, Indian reserves, settlements and Inuit communities) are available on the Statistics Canada website: www12.statcan.gc.ca/nhs-enm/2011/ref/about-apropos/nhs-enm_r012.cfm?Lang=E.

2.3 Questionnaires and instruction guides

The National Household Survey used different forms and questionnaires to collect data from Canadians.

The N1 questionnaire focused on a variety of topics including Aboriginal ancestry, Aboriginal identity, First Nation/Indian band membership and Registered or Treaty Indian status. It was used to enumerate selected private dwellings in mail-back areas.

The N2 questionnaire, the Northern and Reserves Questionnaire, is similar to questionnaire N1. The N2 questionnaire was used to enumerate Indian reserves and settlements, Métis settlements, and Inuit and remote communities. In 2011, there were some areas where there were non-Aboriginal households on leased Aboriginal land and these did not receive the N2 questionnaire. The questions asked on the N2 questionnaire are the same as on the N1 questionnaire, but the examples provided for write-in responses include more Aboriginal groups listed in the ethnic origin/ancestry question, and industries or occupations more commonly found in the North. Examples for the education questions are also geared to the Aboriginal population living in these areas. As well, in the N2 questionnaire, persons living on Indian reserves or settlements were instructed to skip questions 10, 11 and 12 on citizenship, landed immigrant status, and year of immigration, respectively. Refer to www12.statcan.gc.ca/nhs-enm/2011/ref/about-apropos/ques_guide-eng.cfm for more information on the questionnaires and guides.

2.3.1 Questionnaire languages

Translations of the census and NHS questions were available in a number of Aboriginal and immigrant languages, as well as English and French, so that respondents could read the questions in their own language. The 2011 Census and NHS questionnaires were translated into the following eleven Aboriginal languages:

Plains Cree
Swampy Cree
Cree (Quebec)
Dene (Dogrib)
Inuktitut (Labrador)
Inuktitut (Nunavik)
Inuktitut (Nunavut)
Montagnais
Naskapi
Ojibway
Oji-Cree

3 Data processing

The processing phase of the 2011 Census and National Household Survey began with the process of translating responses into meaningful data. This part of the survey cycle was divided into six main activities:

- Receipt, registration and storage
- Imaging and data capture from paper questionnaires
- Edits and failed edit follow-up
- Coding
- Edit and imputation
- Weighting

3.1 Receipt, registration and storage

Statistics Canada's Data Operations Centre (DOC) was the central reception, registration and storage point for electronic and printed questionnaires. Electronic questionnaires were transmitted directly to the DOC's servers, and printed questionnaires were scanned and stored as images. After the quality of the image was confirmed, the data were captured by optical mark recognition (OMR) and intelligent character recognition (ICR). If the image quality was inadequate, the data were captured manually by an operator.

3.2 Imaging and data capture from paper questionnaires

Upon reception and registration of paper questionnaires, the documents went through the following process:

- **Document preparation** – Mailed-back questionnaires were removed from envelopes, and foreign objects such as clips and staples detached. The questionnaires were then organized in batches, by form type, to have their spines cut off (Booklet format forms, only) in preparation for scanning.

- **Scanning** – Seven high-speed scanners were used to create images of each page of each questionnaire.

- **Automated image quality assurance** – An automated system verified the quality of the scanning. Images failing this process were flagged for rescanning or sent to keying.

- **Automated data capture** – Optical mark recognition and optical character recognition technologies were used to extract respondents' data from the images. Where the systems could not recognize the handwriting with sufficient accuracy, data recognition was completed by a census (key) operator.

- **Check-out** – As soon as the questionnaires were processed successfully through all of the above steps, the paper questionnaires were checked out of the system. Check-out is a quality assurance process that ensured the images and captured data are of sufficient quality that the paper questionnaires were no longer required for subsequent processing. Questionnaires that had been flagged as containing errors were pulled at check-out and reprocessed as required.

3.3 Edits and failed edit follow-up

At this stage, a number of automated coverage edits were performed on the respondent data. If multiple questionnaires were received for one household, they were also verified at this stage to determine if they were duplicates (e.g., a husband completed the Internet version and his wife filled in the paper form and mailed it back). Data from questionnaires that failed the edits were forwarded to a processing clerk for verification against the image if available (online questionnaires would not have an image).

Once coverage edits were completed, the household data were subjected to automated completion edits that simulated those that enumerators would have done manually in censuses prior to 2006. They checked for completeness of the responses as well as coverage (e.g., the number of persons in the household). A score was attributed in cases where:

- the processing clerk was not able to resolve a coverage error in the Data Operations Centre
- there was an indication that the respondent was unsure of whether or not (a) person(s) should be included in the household
- there were data indicating that this was a dwelling occupied solely by temporary or foreign residents
- there were many missing or invalid responses.

Although census data were transmitted to one of five Statistics Canada regional offices for failed edit follow-up, NHS data were only subject to follow up where there was total non-response, except in the case of the canvasser areas, where questionnaires were completed during enumeration.

While automated edits were applied to all form types as described above, follow-up was not performed on questionnaires in canvasser areas as this was done during enumeration.

3.4 Coding

The N1 and N2 questionnaires contained questions where answers could be checked off against a list, as well as questions requiring a written response from the respondent in the boxes provided. These written responses (write-in responses) underwent automated coding to assign each response a numeric code, using Statistics Canada reference files, code sets and standard classifications. When the system was unable to assign a code to a particular response, the response was coded manually by a specially trained coder. In 2011, coding was applied to the following variables: relationship to Person 1, place of birth, citizenship, non-official languages, home language, mother tongue, ethnic origin, population group, First Nation/Indian band membership, religion, place of residence 1 year ago, place of residence 5 years ago, place of birth of parents, major field of study, location of study, language at work, industry, occupation and place of work.

3.4.1 Coding of the First Nation/Indian band membership write-in question

Write-in responses to the First Nation/Indian band membership question were coded to a list of over 600 Indian bands. The proportion of responses done by automated coding was 68%. The remaining responses were coded using interactive applications designed specifically for First Nation/Indian band coding. The systems included several reference files such as a file containing different spellings of Indian band names and the corresponding codes, and a file containing geographic codes for Indian reserves, names of Indian reserves, and names of Indian bands that are affiliated with these reserves.[8] The First Nation/Indian band membership data are not available on the dissemination data file but are available on request.

3.5 Edit and imputation

After data capture, and initial editing and coding operations were completed, the data were processed up to the final edit and imputation stage. The final editing detected invalid responses and inconsistencies. This editing was based on rules determined by Statistics Canada's subject-matter analysts. Unanswered questions were also identified. Imputation replaced these missing, invalid or inconsistent responses with plausible values. When carried out properly, imputation can improve data quality by replacing non-responses with plausible responses similar to the ones that the respondents would have given if they had answered the questions. It also has the advantage of producing a complete data set.

The nearest-neighbour-donor method was used to impute NHS data. This method is widely used in the treatment of item non-response. It replaces missing, invalid or inconsistent information about one respondent with values from another 'similar' respondent. The rules for identifying the respondent most similar to the non-respondent may vary with the variables to be imputed. Donor-imputation methods have good properties and generally will not alter the distribution of the data, a drawback of many other imputation techniques. Following nearest-neighbour imputation, consistency of data is assured (see the NHS User Guide Chapter 4 – Data processing).

3.6 Edit and imputation of Aboriginal variables

The edit and imputation of the ethnocultural variables, and specifically the Aboriginal variables, was almost entirely redesigned for 2011, with the primary goal being to streamline the processes and to use, as much as possible, one donor to impute data for a respondent who had provided incomplete or invalid responses on his/her NHS questionnaire.

In 2011, the variables of immigration, citizenship, place of birth, ethnic origin/Aboriginal ancestry, population group/visible minority, Aboriginal group, Registered or Treaty Indian status, and First Nation/Indian band membership were processed together, with the interrelations between these variables clearly defined in advance. Donor imputation for missing information within these variables was done with one donor for all variables, as much as possible.

In 2011, all people requiring imputation, who were not census family children, used a single donor, who was also not a census family child. Census family children who required imputation used a donor within their own census family (sibling or parent). As a result, the imputed records were internally consistent and based on actual full responses, rather than multiple-donor responses that might have donated inconsistent information. This is a definite improvement

8. Statistics Canada acknowledges the expertise and assistance provided by Aboriginal Affairs and Northern Development Canada in the coding of the First Nation/Indian band responses for the 2011 NHS.

over past methods, where units were stratified for imputation based roughly on language and geography, but not on the host of variables that were used in 2011.

The low rates for item non-response and invalid response rates, and the corresponding low imputation rates for Aboriginal variables (Aboriginal group, Registered or Treaty Indian status, and Membership in a First Nation/Indian band) (see Table 2), had little overall impact on data quality.

The 2011 NHS total imputation rates for questions 18 (Aboriginal group), 20 (Registered or Treaty Indian status) and 21 (Membership in a First Nation/Indian band) are shown in Table 2.

Table 2 Imputation rates for Aboriginal group, Registered or Treaty Indian status, and Membership in a First Nation/Indian band, Canada, provinces and territories, 2011 NHS

Provinces and territories	Aboriginal group (%)	Registered or Treaty Indian status (%)	Membership in a First Nation/Indian band (%)
Canada	**3.7**	**4.7**	**3.8**
Newfoundland and Labrador	4.4	6.0	4.2
Prince Edward Island	3.9	6.1	3.9
Nova Scotia	3.7	5.1	3.4
New Brunswick	3.4	4.4	3.2
Quebec	3.6	4.3	3.4
Ontario	4.1	5.1	4.2
Manitoba	3.3	4.3	3.5
Saskatchewan	3.0	4.3	3.5
Alberta	3.4	4.6	3.9
British Columbia	3.5	4.6	3.9
Yukon	3.6	4.9	5.1
Northwest Territories	1.1	1.0	3.4
Nunavut	1.4	1.1	3.4

Source: Statistics Canada, National Household Survey, 2011.

3.7 Weighting

The final responses were weighted so that the data from the sample accurately represent the NHS's target population. The NHS weighting process involved calculating sampling weights, adjusting the weights for the survey's total non-response and calibrating the weights against census totals.

The sampling fraction varied with the questionnaire delivery mode. For the mail delivery mode, about 3 in 10 households (29%) received a questionnaire. For the enumerator delivery mode, the sampling fraction is 1 in 3 households (33%). However, in cases where it was necessary to reach households in remote areas or on Indian reserves or settlements, or Inuit communities where only the interview response mode was offered, no sampling was done and **all** households were invited to participate in the NHS.

Then the sampling weights were adjusted to reflect the targeted non-response follow-up that was done on a subsample of those households that had not responded to the NHS by mid-July 2011.

Subsequent to non-response follow-up, the resulting weight was adjusted for the survey's residual non-response within the subsample. This was done by transferring the weights of non-respondent households to the nearest-neighbour respondent households in the subsample.

Lastly, the weights were calibrated against census population totals for geographic areas known as calibration areas. Weight calibration was performed so that the estimates for an NHS calibration area would be approximately equal to the census counts for that area, for a set of about 60 characteristics common to the NHS and the census. Calibration is a realignment of survey estimates to known population control totals by a minimal modification of the weights. In this case, the census provides a number of counts for various demographic, social and geographic characteristics of the population. These are used as the population controls. The sample survey weights from the NHS are adjusted so that the estimates from the survey match these known counts of the census. With the assurance that their known population composition is maintained, the resulting calibrated weights are then applied to all other variables and characteristics of the survey.

Nevertheless, there may be differences between the NHS estimates and the census counts for common characteristics. Certain factors come into play, that explain these differences, in particular the size of the geographic area and the level of non-response. The smaller the population count in a certain geographic area, the greater the risk that the NHS estimates will be different from the census counts. This issue was present with the 2006 Census long form, but it was less common because of the higher response rates and the calibration method used retained the demographic characteristics for both small and large municipalities. As a guideline to users, whenever the NHS population estimate or its distribution is not similar to the census comparable counts it may indicate quality issues due to non-response. It is suggested that such geographies should be collapsed to a higher level of dissemination. Estimates for Indian reserves that demonstrate a similar effect should be combined with other Indian reserves associated with the same Indian band as the estimates are more likely to yield more reliable estimates.

For additional information on the methodology of the NHS, refer to the *National Household Survey User Guide*, Catalogue no. 99-001-X2011001. All efforts are made to reduce errors in estimation and the Census of Population plays a major role in ensuring the reliability in the estimates of the NHS.

4 Data dissemination

Data on Aboriginal peoples from the 2011 NHS, along with immigration and ethnocultural diversity data were disseminated on May 8, 2011.

In addition, the Education and Labour release (June 26, 2013) and the Income and Housing release (September 11, 2013) included information on Aboriginal peoples, in data tables and in an article in the *NHS in Brief* series (see details below).

4.1 Analytical products

4.1.1 Analytical documents

The analytical document from the initial release, *Aboriginal Peoples in Canada: First Nations People, Métis and Inuit*, Catalogue no. 99-011-X2011001, provides a national overview of key findings from the 2011 National Household Survey for the Aboriginal population. The report is divided into three parts: Aboriginal people – Diverse groups living across the country; The Aboriginal population is young; and Living arrangements of Aboriginal children.

The initial release article was accompanied by a companion analytical article in the *NHS in Brief* series entitled Aboriginal peoples and language, Catalogue no. 99-011-X2011003. As well, as part of the second release of the 2011 NHS data, there is another analytical article in the *NHS in Brief* series entitled The educational attainment of Aboriginal peoples in Canada, Catalogue no. 99-012-X2011003.

In addition to these NHS analytical documents, there is an analytical document on Aboriginal languages, from the 2011 *Census in Brief* series entitled *Aboriginal languages in Canada*, Catalogue no. 98-314-X2011003.

4.2 Data products

4.2.1 NHS Profiles

The *NHS Aboriginal Population Profile* presents information on the Aboriginal identity population from the 2011 National Household Survey (NHS). Data are provided for selected standard geographic areas including Canada, provinces and territories, census divisions, census subdivisions (including Indian reserves, Indian settlements and Inuit communities), census metropolitan areas and census agglomerations, as well as geographic areas such as Indian band areas, Inuit regions and Métis settlements associated with Aboriginal peoples (also known as Aboriginal geographies).

The *NHS Profile*, Catalogue no. 99-004-X presents information from the 2011 National Household Survey (NHS) for various levels of geography, including provinces and territories, census metropolitan areas/census agglomerations, census divisions, census subdivisions, census tracts and federal electoral districts.

In order to provide a comprehensive overview of an area, this product presents data from both the NHS and the 2011 Census. NHS data topics include: Immigration and Ethnocultural Diversity; Aboriginal Peoples; Education and Labour; Mobility and Migration; Income and Housing. 2011 Census data topics include: Population and dwelling counts; Age and sex; Families, households and marital status; Structural type of dwelling and collectives; and Language.

The *Profile for the NHS Special Collection for 13 Indian reserves and Indian settlements in Northern Ontario* provides information from a special data collection following the 2011 National Household Survey (NHS). The census and NHS were not completed during the regular collection period due to forest fires in some communities of Northern Ontario. The NHS Special Collection for 13 Indian reserves and Indian settlements in Northern Ontario (October 2011) was a voluntary survey in which all households in these 13 areas received the questionnaire used to enumerate canvasser areas in the 2011 NHS.

4.2.2 NHS Data Tables – Aboriginal Peoples

The NHS Data Tables – Aboriginal Peoples is a set of tables that provide statistical information about Aboriginal peoples in Canada. The data are presented at the following geographic levels: Canada, provinces and territories; and census metropolitan areas (CMAs) and census agglomerations (CAs).

In addition to these NHS data tables there are also three **2011 Census** tables with data on Aboriginal languages.

Tables:

98-314-XWE2011047 (Languages Spoken Most Often at Home – Detailed Aboriginal Languages (85), Other Languages Spoken Regularly at Home – Aboriginal Languages (10), Mother Tongue – Detailed Aboriginal Languages (85) ...)
www5.statcan.gc.ca/bsolc/olc-cel/olc-cel?catno=98-314-X2011047&lang=eng

98-314-XWE2011048 (Mother Tongue – Detailed Aboriginal Languages (85), Languages Spoken Most Often at Home – Detailed Aboriginal Languages (85), Other Languages Spoken Regularly at Home – Aboriginal Languages (12), ...)
www5.statcan.gc.ca/bsolc/olc-cel/olc-cel?catno=98-314-X2011048&lang=eng

and

98-314-XWE2011050 (Mother Tongue – Detailed Aboriginal Languages (79), Single and Multiple Language Responses (3), ...)
www5.statcan.gc.ca/bsolc/olc-cel/olc-cel?catno=98-314-X2011050&lang=eng.

4.2.3 NHS Focus on Geography Series

The *NHS Focus on Geography Series*, Catalogue no. 99-010-X, provides a quick access to key results from the NHS for different levels of geography and topics. These topics include: Aboriginal Peoples, Immigration and Ethno-cultural Diversity, Education, Labour, Income and Housing. Data are presented for the population living in private households and/or for the number of occupied private dwellings.

4.3 Reference products

4.3.1 NHS Dictionary

The *National Household Survey Dictionary*, Catalogue no. 99-000-X, is a reference document which contains detailed definitions of concepts, universes, variables, and geographic terms used in the NHS. By referring to the NHS Dictionary, both beginner and intermediate data users will gain a better understanding of the data.

4.3.2 Aboriginal Peoples Reference Guide

The *Aboriginal Peoples Reference Guide, National Household Survey*, Catalogue no. 99-011-X2011006, provides information that enables users to effectively use, apply and interpret data from the 2011 NHS. This guide contains definitions and explanations of concepts, classifications, data quality and comparability to other sources. Additional information is included for specific variables to help general users better understand the concepts and questions used in the NHS.

4.3.3 Incompletely enumerated Indian reserves and Indian settlements

The list entitled *Incompletely enumerated Indian reserves and Indian settlements in alphabetical order* provide the name, type, province, census division, census metropolitan area or census agglomeration, and the enumeration status (reasons for absence of data) for the 36 incompletely enumerated Indian reserves and settlements in the 2011 NHS.

4.3.4 Custom products

Statistics Canada also offers custom tabulation services, on a cost-recovery basis, to data users who cannot find published NHS data that respond to their needs. To request custom tabulations or other services, external clients may call 1-800-263-1136 or email infostats@statcan.gc.ca.

5 Data quality assessment and indicators

The objective of data quality assessment is to evaluate the overall quality of survey data, so as to improve our understanding of how and where errors occur, and to inform users of the reliability of the data.

Although there are several potential sources of error, they can be grouped into two types: sampling error and non-sampling error. The former is present because when we estimate a characteristic, we are measuring only part of the population that may not represent the whole population, particularly with high non-response. The latter covers all errors that are not related to sampling (coverage, response, and processing errors).

This section is divided into three main sub-sections: the first deals with different sources of error; the second looks at data suppression related to confidentiality and data quality; and the third deals with coverage error related to Indian reserve communities.

5.1 Sources of error

5.1.1 Sampling error

The objective of the NHS is to produce estimates from a number of questions for a wide variety of geographies, ranging from very large areas (such as provinces and census metropolitan areas) to very small areas (such as neighbourhoods and municipalities), and for various population groups such as Aboriginal peoples and immigrants. These groups also vary in size, especially when cross-classified by geographic area. Such groupings are generally referred to as 'domains of interest.'

With a sampling rate of about 3 in 10 and a response rate of 68.6%, it is estimated that about 21% of the Canadian population participated in the NHS. Nevertheless, given the voluntary nature of the NHS, the quality of domain estimates may vary appreciably, in particular because of the variation in response rates from domain to domain.[9] Errors from non-response will have an effect on the measurement of sampling error, as it introduces non-response variability in the estimate.

The sampling error of the estimate is often expressed as a coefficient of variation (CV) which is the ratio of the standard error of the estimate over the estimate itself, expressed as a percentage. The CV is used to give an indication of the uncertainty associated with the estimates. While the CVs measured in the NHS are in the same magnitude as those from the 2006 Census long form, this may not be the case for lower levels of geography (see Note: NHS coefficients of variation for a comparison of estimates at various levels of geography for the 2011 NHS and the 2006 Census long form). A known effect of non-response is the reduction on the effective sample yield for any given area. The sampling rate for the NHS was purposefully much higher than the rate for the 2006 Census long form to compensate for the expected rate of non-response. Overall, this strategy worked as expected, however for lower areas of geography or for specific population domains, the differential response rates will still have an effect on the precision of the estimate. As non-response increases, the number of sample cases with a response are reduced, thus increasing the standard error and the CV.

Coefficients of variation are available for selected variables for Canada, provinces and territories. Downloadable files (CSV [87 kb] | TAB [77 kb]) are also available.

---

9. For more information on sampling error, refer to the National Household Survey User Guide, Catalogue no. 99-001-X2011001, www12.statcan.gc.ca/nhs-enm/2011/ref/nhs-enm_guide/guide_4-eng.cfm#A_5_1.

5.1.2 Non-sampling error

Besides sampling, a number of factors can cause errors in a survey's results. Respondents may be missed, incorrectly enumerated or counted more than once (**coverage errors**). Respondents or sometimes the survey representative, misunderstands a question and records an incorrect response or simply uses the wrong response box (**response errors**), and responses may be entered incorrectly during data capture coding and processing (**processing errors**).[10] These are examples of non-sampling errors that were thoroughly accounted for at every stage of collection and processing to mitigate their impact.

In addition, in every voluntary survey, non-response can also introduce error to the estimates from the survey's variables. A substantial portion of the non-sampling error (**measurement errors**) can be associated with non-response. There is a distinction to be made between partial non-response (lack of response to one or some questions) and total non-response (lack of response to the survey because the household could not be reached or refused to participate).

Total non-response to the survey not only reduces the effective number of participants to the survey (effect on **sampling error**), it is also likely to bias the estimates from the survey, because non-respondents tend to have different characteristics than respondents. As a result, there is a risk that the results will not be representative of the actual population.

There is non-response bias when a survey's non-respondents differ from its respondents in how they would have responded to the information collected. In that case, the higher a survey's non-response is, the greater the risk of non-response bias. Non-response weight adjustments are meant to compensate for these errors. The quality of the estimates can be affected if such a bias persists.

Several methods can be used during data collection or processing to minimize non-response bias. NHS non-response follow-up was planned in such a way as to maximize the survey's response rate and control potential non-response bias due to the survey's voluntary nature. Ultimately, non-response adjustments done to the survey weights resulting from the non-response follow-up and calibration to known demographic and geographic benchmarks reduce bias significantly.

The NHS has an unweighted response rate of 68.6%, and a weighted response rate of 77.2%.

The weighted response rate is the measurement of the reduced impact of non-response after non-response follow-up. Statistics Canada conducted several studies and various simulations, before and after collection, to assess the risk and extent of the potential bias. A number of measures were taken to mitigate its effects.[11]

---

10. Processing errors can occur at various steps including coding, when 'write-in responses' are transformed into numerical codes; data capture, when responses are transferred from the questionnaire in an electronic format, by optical character recognition methods or key-entry operators; and imputation, when a 'valid,' but not necessarily correct, response is inserted into a record by the computer to replace missing or 'invalid' data ('valid' and 'invalid' referring to whether or not the response is consistent with other information on the record).
11. For more information on the NHS data quality assessment process and indicators, refer to NHS data quality assessment process and indicators.

5.1.2.1 Non-response bias for Aboriginal variables

Several data sources were used to evaluate the NHS estimates for Aboriginal variables such as: 2011 Census results for mother tongue (since a relationship exists between language and Aboriginal identity, Registered or Treaty Indian status and Membership in a First Nation/Indian band); the 2001 and 2006 censuses; Population Projections by Aboriginal Identity in Canada; and administrative data pertaining to Registered Indians from Aboriginal Affairs and Northern Development Canada (AANDC).

It is impossible to definitively determine how much the NHS may be affected by non-response bias. However, based on information from other data sources, evidence of non-response bias does exist for certain populations and for certain geographic areas. Generally, the risk of bias increases for lower levels of geography and for smaller populations.

On the basis of the estimates and trends from the sources mentioned above, evidence suggests that biases were in general well mitigated, with exceptions. The Inuit population living outside of Inuit Nunangat appears to be slightly overestimated at the national level. The magnitude of this overestimation appears to be higher (and more variable) for some smaller geographic areas. Additionally, while the NHS results show increases of the Métis population and the First Nations population living off reserve as compared with data from the 2006 Census, many factors, other than non-response bias, could explain the growth of these populations, including changes in reporting patterns and the propensity of people to self-identify as an Aboriginal person.

5.2 Data suppression related to confidentiality and data quality

Data disseminated by the NHS are also subjected to a variety of automated and manual processes to determine whether the data need to be suppressed in order to maintain confidentiality (non-disclosure) and data quality.

5.2.1 Data suppression related to confidentiality (non-disclosure)

All NHS data are subject to confidentiality suppression rules, to ensure non-disclosure of individual respondent identity and characteristics. The following describes the various suppression rules used to ensure confidentiality.

5.2.1.1 Area suppression for standard geographic areas

Area suppression is used to remove all characteristic data for geographic areas below a specified population size. The specified population size for all standard[12] areas or aggregations of standard areas is 40, except for blocks, block-faces or postal codes. Consequently, no characteristics or tabulated data are released for areas, such as CSDs (municipalities or Indian reserves or settlements and unorganized territories) below a total population size of 40.[13]

5.2.1.2 Area suppression for income characteristics data

Area suppression is used to replace all income characteristic data with an 'x' for geographic areas with populations and/or number of households below a specific threshold.

12. For more information on standard areas, refer to the 2011 *Census Dictionary*, Catalogue no. 98-301-X, www12.statcan.gc.ca/census-recensement/2011/ref/dict/index-eng.cfm.
13. For more information refer to area suppression for NHS standard and non standard geographic areas.

If an NHS tabulation contains quantitative income data (e.g., total income, wages), qualitative data based on income concepts (e.g., low income before tax status) or derived data based on quantitative income variables (e.g., indexes) for individuals, families or households, then the following rule applies: income characteristic data are replaced with an 'x' for areas where the estimated population is less than 250 or where the number of private households is less than 40. The private household threshold does not apply for tabulations based on place of work geographies.

5.2.1.3 Random rounding

All estimates in NHS tabulations are subjected to a process called random rounding. Random rounding is a method used to modify an estimate to a value ending in '5' or '0.' It is either greater or less than the original value. This reduces the possibility of identifying individuals within the tabulations.

All counts greater than 10 are rounded to base 5, counts less than 10 are rounded to base 10. This means that any counts less than 10 will always be changed to 0 or 10. Table 3 below shows the effect of rounding on counts with a value less than 10.

Table 3 Random rounding frequency

Count of	Will round to 0	Will round to 10
1	9 times out of 10	1 time out of 10
2	8 times out of 10	2 times out of 10
3	7 times out of 10	3 times out of 10
4	6 times out of 10	4 times out of 10
5	5 times out of 10	5 times out of 10
6	4 times out of 10	6 times out of 10
7	3 times out of 10	7 times out of 10
8	2 times out of 10	8 times out of 10
9	1 time out of 10	9 times out of 10
0	Always	Never

The random rounding algorithm uses a random seed value to initiate the rounding pattern for tables. In these routines, the method used to seed the pattern can result in the same count in the same table being rounded up in one execution and rounded down in the next.

5.2.1.4 Suppression of NHS estimates for confidentiality reasons

The previous section discussed random rounding for estimates in NHS tabulations and minimum population thresholds in order to protect the anonymity of respondents. Random rounding is used as a means of protecting confidentiality in counts when characteristics become rare. Analysis of NHS estimates revealed that weighted data may result in high estimates that meet the aforementioned population suppression threshold; in these cases individuals with rare characteristics could be more easily identified in a table, particularly if their characteristics are publicly known.

Consequently, for all quantitative variables, a statistic is suppressed if the number of actual records used in the calculation (not rounded or weighted) is less than 4. For quantile statistics, an alternate minimum number of records apply: for quartiles, quintiles and deciles, 20 records are required, and for percentiles, 400 records are required.

For more information on confidentiality (non-disclosure) rules, refer to the *Data Quality and Confidentiality Standards and Guidelines (Public)* (www12.statcan.gc.ca/nhs-enm/2011/ref/DQ-QD/index-eng.cfm).

5.2.2 Data quality indicators

Data dissemination, in addition to being suppressed (limited) for confidentiality (non-disclosure) reasons, may also be limited as a result of unacceptable data quality (which will subsequently be referred to as **data quality**).[14]

5.2.2.1 Global non-response rates

The global non-response rate (GNR) is an indicator of data quality which combines complete non-response and partial non-response to the survey. A smaller GNR indicates a lower risk of non-response bias, i.e., a lower risk of lack of accuracy. Global non-response rates are determined for each of the NHS geographic areas. These areas are flagged on the database according to the non-response rate. Geographic areas with a global non-response rate higher than or equal to 50% are suppressed from standard data products but will be available as a custom request. Geographic areas with a global non-response rate lower than 50% are identified in tabulations, but not suppressed.[15] In addition, while characteristics data were suppressed for these areas, they were included in all higher geographic level tabulations.

Canada has a total of 147 census metropolitan areas (CMAs) and census agglomerations (CAs). For all of these areas, the global non-response rate is less than 50% and published NHS data are available in standard products. In addition, NHS standard products are available for all 293 census divisions (CDs) and all 308 federal electoral districts (FEDs).

With a global non-response rate threshold of 50% for the release of NHS data, estimates are published for a majority of the total 5,253 census subdivisions (CSDs) or municipalities. A total of 4,567 CSDs have an estimated population of more than 40 (for confidentiality reasons, those with a population of less than 40 are not published); 686 CSDs have a population of less than 40 (including zero population – or uninhabited CSDs). Of the 4,567 CSDs with a population of more than 40, NHS estimates are available in standard products for 3,439 (75.3%).[16]

5.2.3 Other occurrences when data are suppressed or not available

In addition to being suppressed for confidentiality and data quality reasons, data may also be suppressed or not available for reasons related to data collection.

5.2.3.1 Suppression of citizenship, landed immigrant status and period of immigration data – Indian reserve N2 suppression

Suppression of data also occurs when certain questions are not asked of all respondents. Persons living on Indian reserves and Indian settlements who were enumerated with the 2011 NHS N2 questionnaire were not asked the questions on citizenship (Question 10), landed immigrant status (Question 11) and year of immigration (Question 12). However, it was possible that a census subdivision (CSD) or lower geographic area was enumerated using both the N2 questionnaire (for the on-reserve population) and the N1 questionnaire (for the off-reserve population). In this case,

---

14. The NHS standard products contain the following two confidentiality and data quality symbols:
... not applicable
x suppressed to meet the confidentiality requirements of the *Statistics Act*
15. For more information concerning global non-response rates in the NHS, refer to the *Data Quality and Confidentiality Standards and Guidelines (Public)* (www12.statcan.gc.ca/nhs-enm/2011/ref/DQ-QD/guide_4-eng.cfm#A_4_1_1_1_3).
16. Refer to the National Household Survey User Guide for provincial/territorial distribution of published CSDs www12.statcan.ca/nhs-enm/2011/ref/nhs-enm_guide/guide_5-eng.cfm#A_6_4.

the following rules were used to determine if suppression had to be applied to all citizenship and immigration data for that CSD (or lower geographic area):

1. If the population estimate from N1 questionnaires was higher than the population estimate from N2 questionnaires (based on weighted results), then citizenship and immigration estimates were included in the CSD estimates.

2. If the population estimate from N2 questionnaires was higher than or equal to the population estimate from N1 questionnaires (based on weighted results), then citizenship and immigration estimates were excluded from the CSD estimates.

Consequently, citizenship, landed immigrant status and period of immigration data are suppressed for Indian reserves and Indian settlements at census subdivision and lower levels of geography where the majority of the population was enumerated with the N2 questionnaire. These data are, however, included in the totals for larger geographic areas, such as census divisions and provinces.

For a complete list of Indian reserves and Indian settlements for which citizenship, landed immigrant status and period of immigration data are suppressed, refer to www12.statcan.gc.ca/nhs-enm/2011/ref/sup_N2-eng.cfm.

5.2.3.2 Incompletely enumerated areas

In 2011, there were a total of 36 Indian reserves and Indian settlements reported[17] as 'incompletely enumerated' in the NHS. For 18 reserves or settlements, census enumeration was either not permitted or was interrupted before it could be completed and so the NHS was not administered in those areas. For four reserves or settlements, census enumeration was completed however data collection for the NHS was not permitted or interrupted, and in one reserve or settlement it was determined that there was no resident population contrary to what was erroneously reported in the census. In the case of 13 reserves in Northern Ontario, enumeration was delayed because of natural events (specifically forest fires) and estimates for these communities are not included in geographic areas that include these communities (e.g., provincial and national estimates). For these 13 reserves, separate tables are made available (see Section 4.2.1).

5.2.4 Data availability from the NHS for census subdivisions

5.2.4.1 Data availability from the NHS for communities (census subdivisions) with Aboriginal identity population

Of the total 5,253 CSDs in Canada, 3,972 CSDs have estimates of Aboriginal identity population. Table 4 shows the number of CSDs for which NHS Aboriginal identity estimates are available, in addition to the number of CSDs for which Aboriginal identity estimates are suppressed.

Among the 3,972 CSDs that have Aboriginal identity population, NHS Aboriginal identity estimates are available in 2,385 (60.0%). Aboriginal identity estimates are suppressed for confidentiality reasons in a total of 1,044 CSDs (in 217 because the total population is less than 40, and in 827 because Aboriginal identity cell counts are less than 4 – suppression is applied for confidentiality reasons). In addition, Aboriginal identity estimates (as well as other characteristics) are suppressed for 543 CSDs due to data quality – the global non-response rate is 50% or greater – in these CSDs data are of unacceptable quality, the response rates are not high enough to produce a valid statistical picture.

17. For more information on 2011 NHS incompletely enumerated Indian reserves and Indian settlements, refer to www12.statcan.gc.ca/nhs-enm/2011/ref/aboriginal-autochtones-eng.cfm.

Table 4 CSDs with Aboriginal identity estimates, by type of data suppression and data availability, Canada, provinces and territories, 2011 NHS

	Number of CSDs with Aboriginal identity estimates													
	Canada	N.L.	P.E.I.	N.S.	N.B.	Que.	Ont.	Man.	Sask.	Alta.	B.C.	Y.T.	N.W.T.	Nvt.
Total CSDs	5,253	376	113	99	273	1,285	574	287	959	435	743	37	41	31
Total incompletely enumerated reserves (CSDs)	36	0	0	0	0	7	22	3	2	1	1	0	0	0
Total CSDs with no population	316	10	0	4	3	102	16	10	60	20	76	6	3	6
Total CSDs (with population)	4,901	366	113	95	270	1,176	536	274	897	414	666	31	38	25
CSDs with Aboriginal identity population	3,972	206	72	93	221	877	522	268	616	346	658	30	38	25
CSDs with Aboriginal identity population **and** total population less than 40 **(for confidentiality reasons, CSDs with a population of less than 40 are not published)**	217	1	2	6	2	1	4	5	42	10	134	6	4	0
CSDs with Aboriginal identity population **and** total population 40 or greater **and** Aboriginal identity cell count less than 4 **(for these CSDs NHS counts are suppressed for confidentiality reasons)**	827	67	44	8	77	310	27	29	199	49	17	0	0	0
CSDs with Aboriginal identity population **and** total population 40 or greater **and** Aboriginal identity cell count greater than 3 **and** GNR equal to or greater than 50% **(for these CSDs, NHS counts are suppressed for data quality reasons)**	543	29	7	9	24	57	83	66	121	50	84	9	0	4
CSDs for which Aboriginal identity estimates are available – CSDs with Aboriginal identity population **and** total population 40 or greater **and** Aboriginal identity cell count greater than 3 **and** GNR is less than 50%	2,385	109	19	70	118	509	408	168	254	237	423	15	34	21

Source: Statistics Canada, National Household Survey, 2011.

The proportion of CSDs for which Aboriginal identity estimates are available (including population estimates and characteristics) varies by province and territory. With the exception of Prince Edward Island (26.4%) and Saskatchewan (41.2%), in all other provinces and territories, estimates are available for at least half of the CSDs with Aboriginal identity estimates.

Methodological changes

The availability of Aboriginal estimates, at the CSD level, reflects a change made to the data quality threshold over the 2006 Census release. The NHS 50% threshold is based on studies of the global non-response rate in relation to the indicators of non-response bias (see Section 5.5 of the *National Household Survey User Guide*, Catalogue no. 99-001-X2011001). The studies showed that with a global non-response rate of 50% or more, the bias was so large that the estimates were not of sufficiently high quality.

5.2.4.2 Data availability from the NHS for 'on reserve' communities (CSDs)

The 2011 NHS and the 2011 Census 'on reserve' area of residence is comprised of a total of 997 CSDs, including 159 uninhabited CSDs, 36 incompletely enumerated Indian reserves and settlements and 802 habited CSDs. Among the 802 'on reserve' inhabited CSDs (with population), three (3) CSDs did not have an Aboriginal identity population.[18] Of the 799 CSDs with an Aboriginal identity population, 190 were suppressed because they had less than 40 total population (Aboriginal and non-Aboriginal), one additional reserve CSD was suppressed because the Aboriginal identity cell count was less than four (4), and another 36 CSDs were suppressed for data quality reasons – the GNR was greater than or equal to 50%. Overall, Aboriginal data (population estimates and characteristics) are available for 572 'on reserve' CSDs. Table 5 shows the number of communities (CSDs) defined as 'on reserve' for which NHS Aboriginal identity data are available.

Table 5 'On reserve' CSDs, by type of data suppression and data availability, Canada, provinces and territories, 2011 NHS

	Number of 'on reserve' CSDs with Aboriginal identity estimates													
	Canada	N.L.	P.E.I.	N.S.	N.B.	Que.	Ont.	Man.	Sask.	Alta.	B.C.	Y.T.	N.W.T.	Nvt.
Total CSDs	997	3	4	25	18	42	144	79	170	85	425	0	2	0
Total incompletely enumerated reserves (CSDs)	36	0	0	0	0	7	22	3	2	1	1	0	0	0
Total CSDs with no population	159	0	0	4	0	0	12	6	52	11	74	0	0	0
Total CSDs (with population)	802	3	4	21	18	35	110	70	116	73	350	0	2	0
CSDs with Aboriginal identity population	799	3	4	21	18	35	109	70	116	73	348	0	2	0
CSDs with Aboriginal identity population **and** total population less than 40 **(for confidentiality reasons, CSDs with a population of less than 40 are not published)**	190	0	2	6	2	1	3	5	27	9	134	0	1	0
CSDs with Aboriginal identity population **and** total population 40 or greater **and** Aboriginal identity cell count less than 4 **(for these CSDs, NHS counts are suppressed for confidentiality reasons)**	1	0	0	0	0	0	0	0	0	0	1	0	0	0
CSDs with Aboriginal identity population **and** total population 40 or greater and Aboriginal identity cell count greater than 3 and GNR equal to or greater than 50% **(for these CSDs, NHS counts are suppressed for data quality reasons)**	36	0	0	0	0	1	16	4	7	0	8	0	0	0
CSDs for which Aboriginal identity estimates are available – CSDs with Aboriginal identity population **and** total population 40 or greater **and** Aboriginal identity cell count greater than 3 **and** GNR is less than 50%	572	3	2	15	16	33	90	61	82	64	205	0	1	0

Source: Statistics Canada, National Household Survey, 2011.

The availability of Aboriginal identity data for 'on reserve' communities varies among the provinces and territories. With the exception of Prince Edward Island (50.0%), Northwest Territories (50.0%) and British Columbia (58.9%), in all other provinces and territories, Aboriginal identity data are available in at least 70% of the 'on reserve' CSDs.

18. These may be reserves with lease land occupied by non-Aboriginal persons or they may be reserves where the Aboriginal population participated in the census but did not participate in the NHS. Note that these three reserves have less than 40 total population, and for confidentiality reasons, no characteristics data are released, only total population estimates are published.

Table 6 Data availability for on-reserve communities, 2011 NHS, 2006 and 2001 censuses

	2011 NHS		2006 Census		2001 Census	
	Number	%	Number	%	Number	%
Total reserve communities – inhabited	837	100.0	887	100.0	881	100.0
Incompletely enumerated reserve communities – data not available	35	4.2	22	2.5	30	3.4
Reserve communities where total population size is less than 40	193	23.1	184	20.7	190	21.6
Reserve communities where total population size is equal to or greater than 40 for which partial[1] data are available	37	4.4	106	12.0	64	7.3
Reserve communities with total population size is equal to or greater than 40 for which full[2] data are available	572	68.3	575	64.8	597	67.8

Notes:
1. Partial data refers to population and dwelling counts only – no characteristics data are available.
2. Full data refers to population and dwelling counts as well as characteristics data.
Sources: Statistics Canada, censuses of population, 2001 and 2006, and National Household Survey, 2011.

Changes in the availability of information over time can be linked to changes in methodology which affect the level of global non-response rates, as well as the additional suppression for confidentiality reasons. In the 2011 NHS, information was suppressed when the unweighted and unrounded cell count contributing to the estimates is less than four (4). Another change regarding the amount of data available between the NHS and the previous censuses is the global non-response rate (GNR) which is used to determine what data are released – in the previous censuses, data were suppressed for CSDs with GNR ≥ 25%, whereas in the NHS, data are suppressed for CSDs with GNR ≥ 50%. The 50% threshold for the NHS is based on studies of correlation between the global non-response rate and indicators of non-response bias. The studies showed that the relationship between potential average biases and non-response was generally acceptable up to a GNR threshold of 50%, allowing for the release of estimates of sufficiently high quality.

Overall, when all groupings of data availability are considered, the data availability for 'on reserve' communities (CSDs) have increased in 2011. The proportion of reserves for which the complete set of data (full data) are available has gone from 67.8% in 2001, to 64.8% in 2006 (using among others, data quality threshold of a GNR < 25%) and to 68.3% in the 2011 NHS (using the data quality threshold of a GNR < 50%).

As noted earlier, there were a total of 36 Indian reserves and Indian settlements that were 'incompletely enumerated' in the NHS, compared to 22 in 2006 and 30 in 2001. However, it should be noted that the 36 incompletely enumerated reserves in 2011 include 13 reserves in Northern Ontario for which NHS enumeration was not possible at the time of NHS data collection in May 2011, because of forest fires. These 13 reserves were enumerated later and their data are released in a separate set of special tables (see Section 4.2.1). The proportion of incompletely enumerated reserves for which no data are available remained almost unchanged (2.8% vs. 2.5%) from 2006 to 2011, when the 13 Northern Ontario reserves are excluded from the 2011 estimate.

There was a slight increase from 2006 to 2011 in the proportion of small reserves with a total population size less than 40 for which only population counts are available, from 20.7% in 2006 to 23.1% in 2011.

5.3 Coverage

There are two types of coverage error. **Population undercoverage** refers to the error of excluding someone who should have been enumerated. **Population overcoverage** refers to the error of either enumerating someone more than once or including someone who should not have been enumerated. Undercoverage is more common than overcoverage. The net impact of undercoverage and overcoverage on the size of a population of interest is **population net undercoverage**. Net undercoverage is calculated as the number of persons excluded who should have been enumerated (undercoverage) less the number of excess enumerations of persons enumerated more than once (overcoverage). It is the net of undercoverage and overcoverage, census **net population undercoverage** that quantifies the net number of persons missed by the census.

This section presents estimates of **census net population undercoverage for the 2011 Census** for people residing on participating Indian reserves and settlements including people without Aboriginal identify.

Coverage error generally occurs during the field collection stage of the census. Examples of undercoverage and overcoverage are:

Examples of undercoverage

1. A person temporarily out of the country during the collection of the census is missed.
2. A questionnaire was returned but someone who lived there was not included.
3. The dwelling never received a questionnaire.
4. Persons residing at more than one address may be missed at both addresses because of the uncertainty of what is their main address.
5. Persons who do not reside at a fixed address are often missed by the census.

Examples of overcoverage

1. Children whose parents live in separate households where each parent includes the children in their questionnaire.
2. Young adults, newly away from home perhaps searching for work or attending a post-secondary institution, who are listed there and at home by their parents.
3. Persons whose employment requires them to live away from home. They are listed at both locations.
4. Persons in institutions who are also listed by their families as living at home.

5.3.1 Net undercoverage error for participating reserves

Table 7 gives estimates of 2011 Census net undercoverage **for persons living on participating reserves** for Canada and for each province and territory.[19] The split between persons with Aboriginal identity and persons without Aboriginal identity is not available. The rate of census net undercoverage, on a net basis indicates what proportion of the entire population that should have been enumerated but was missed in 2011 Census tabulations. Negative estimates mean that the estimates of overcoverage were higher than the estimates of undercoverage.

19. The estimates do not include the reserves that did not participate to the 2006 Census, but participated to the 2011 Census.

Table 7 The 2011 Census population net undercoverage for participating reserve communities, Canada, provinces and territories

Provinces and territories	Census count	Census net undercoverage			
	Number	Estimated number	Standard error	Estimated rate (%)	Standard error (%)
Canada	351,394	10,125	6,135	2.80	1.65
Newfoundland Labrador	3,165	379	317	10.68	7.98
Prince Edward Island	514	-6	23	-1.16	4.61
Nova Scotia	9,629	-102	311	-1.08	3.30
New Brunswick	7,855	658	389	7.73	4.21
Quebec	40,455	-1,965	2,381	-5.11	6.50
Ontario	42,328	613	2,938	1.43	6.74
Manitoba	64,000	2,478	1,962	3.73	2.84
Saskatchewan	56,977	-765	2,017	-1.36	3.64
Alberta	47,050	5,815	2,865	11.00	4.82
British Columbia	79,127	2,998	2,625	3.65	3.08
Yukon	0	0	0	0.00	0.00
Northwest Territories	294	23	31	7.11	9.14
Nunavut	0	0	0	0.00	0.00

Source: Statistics Canada, 2011 Census of Population.

The estimate of population overcoverage, for a particular geography such as participating reserves, includes persons who appear on questionnaires for two dwellings where at least one of the dwellings is on reserve. The other dwelling may be on the same reserve, on a different reserve, or not on a reserve. Since the Census Overcoverage Study (COS) does not determine at which dwelling an individual should have been listed at, the assumption is made that it is equally likely that the individual should have been listed at the first dwelling as at the second dwelling. Therefore, in order to produce estimates of overcoverage, half of the weight for the person is assigned to each dwelling. This concept is important for small domains such as the on reserve population. About half of the overcoverage cases involving a dwelling on reserve also involved a dwelling off reserve.

5.3.1.1 Data sources

The estimates of 2011 Census of Population coverage error are derived from 2011 Census data and the results of two studies. The Reverse Record Check (RRC) measures population undercoverage while the Census Overcoverage Study (COS) measures population overcoverage. In the RRC, a random sample of individuals representing the census target population is taken from frames independent of the 2011 Census. The 2011 Census database is searched to determine if these people had indeed been enumerated. When required, an interview (mostly by telephone) was conducted to collect further information to declare the individual as in or not in scope for the census, and when in scope, to provide further data to ascertain that individual's coverage status.

Overcoverage is measured by matching the 2011 Census database to a partial list of persons who should have been enumerated (a list constructed from administrative data sources), and by matching the 2011 Census database to itself. The COS applies statistical matching, which identifies matches that are close or exact. Pairs of potential duplicates are sampled and the sampled person's name and demographic characteristics are used to identify the cases of duplication.

For more information on 2011 Census population coverage error, refer to www.statcan.gc.ca/daily-quotidien/130926/t130926b001-eng.htm.

5.3.2 Coverage error for incompletely enumerated reserves and settlements

As noted earlier, some Indian reserves and settlements did not participate in the census as enumeration was not permitted or was interrupted before completion. In 2011, there were 31 incompletely enumerated reserves in the census (36 in the NHS). For 13 of these reserves, enumeration was prevented by forest fires in Northern Ontario at the time of the census. Census and NHS collection for these 13 reserves was conducted at a later date (fall 2011), and these data are disseminated in a special series of tables (see Section 4.2.1) and are not included in any census or NHS tabulation because of the different collection period. For the remaining 18 incompletely enumerated reserves, census/NHS data are not available and therefore have not been included in any census or NHS tabulation. For four additional reserves, while the census information exists, no NHS data was collected and therefore data for these reserves have not been included in NHS tabulations. Of the five reserves incompletely enumerated in the NHS, only four are considered as participating reserves and their coverage measurements are included in the Section 5.3.1. The remaining reserve, Opaskwayak Cree Nation 27A (Carrot River) was determined to be uninhabited, despite being listed as inhabited in the census.

These areas present unique problems for the coverage studies and for the Population Estimates Program. The survey population of the Reverse Record Check (RRC) does not include those residents where the census was unable to collect any data. However, the Population Estimates Program requires an estimate of the permanent resident population living in these areas. For the 13 reserves in Northern Ontario, a base estimate exists and so do not need to be estimated by a model. However, as neither the census nor the RRC is in a position to produce an estimate of the population living in the remaining 18 areas, a model-based methodology was used for these reserves. The resulting estimates should be used with caution as they are based entirely on a model. Table 8 gives the national model results.

Table 8 Model estimated counts and rates for incompletely enumerated Indian reserves (IER) and settlements for Canada, 2001, 2006 and 2011

	Canada
2001 estimate of IER population	34,992
2006 estimate of IER population	40,115
2011 estimated census count	37,574
2011 net undercoverage rate	-0.4%
2011 net undercoverage	-182
2011 population estimate for IER	37,392

Note: These estimates should be used with caution as they are based on a model whose assumptions cannot be verified.
Sources: Statistics Canada, 2001, 2006 and 2011 censuses of population.

The 2011 net undercoverage rate is different from the rate presented in Section 5.3.1 for participating Indian reserves and settlements as it is calculated by dividing the 2011 net undercoverage for incompletely enumerated Indian reserves and settlements by the corresponding adjusted 2011 population estimate.

In the 2006 Census, 22 reserves, with an estimated 40,100 persons, were classified as 'incompletely enumerated.' Among the 31 reserves and settlements considered as incompletely enumerated in the 2011 Census, 18 were 'incompletely enumerated' or 'refusal' while the other 13 were enumerated at a later date. The 2011 estimates of the incompletely enumerated population are approximately 6.8% lower than the 2006 estimates.

5.3.2.1 Estimation model

A two-step estimation model was developed to estimate the population. The first step uses a simple linear regression to predict the census count in 2011 for the 18 reserves where no data was collected. The linear regression was constructed using all Indian reserves that were completely enumerated in both the 2006 Census and the 2011 Census. The model assumes a linear growth from 2006 to 2011 for all provinces with separate estimates for the intercept and the regression parameters for each province. The model was evaluated for the basic regression assumptions of independence of errors, homogeneity of variances and normality of errors. For the 13 reserves where late enumeration was done, their counts were used as enumerated for this first step.

For each incompletely enumerated reserve, the input variable for the regression model was either the actual census count in 2006 or the best predicted census count from the 2006 model, or the late enumeration for the 13 reserves in Northern Ontario. The output of the model was the estimated census count in 2011.

The second step is done to produce consistency with the results of the census coverage studies. An adjustment was made to the estimated 'census' count to account for net undercoverage of all subjected census counts. Net undercoverage for the incompletely enumerated reserves was estimated by calculating the net undercoverage rate for all completely enumerated reserves in each province and then applying that rate to the estimated 'census' count of all the incompletely enumerated Indian reserves in the province. The estimated 'census' count and the 'estimated net missed persons' in each reserve were then summed to create an 'estimated' population for the incompletely enumerated Indian reserves.

6 Comparability

6.1 Historical comparability

The counts for Aboriginal groups, Registered or Treaty Indian status group and First Nation/Indian band membership may change over time for a number of reasons. Part of the change (increase or decrease) in the number of people reporting Aboriginal group, Registered or Treaty Indian status or First Nation/Indian band membership is due to demographic growth, another part is due to changes in reporting patterns between Aboriginal groups and between Aboriginal and non-Aboriginal populations.

Changes in Aboriginal counts over time can also be affected by differences in the wording and in the format of Aboriginal questions; legislative changes; differences in the set of incompletely enumerated reserves; and changes made to the definition of reserves.

As well, the counts for the Aboriginal variables may also have changed as a result of differences in edit and imputation (E & I) procedures, and changes in methodology between the 2011 NHS and the 2006 and 2001 censuses – for example changes related to weighting and calibration, changes to the universe, etc.

6.1.1 Differences in wording and format of Aboriginal questions

There are various ways to define the Aboriginal population based on the four questions asked in the NHS. These questions are Aboriginal ancestry (Question 17 on ethnic origin); Aboriginal group (Question 18); Registered or Treaty Indian status (Question 20); and Membership in a First Nation/Indian band (Question 21).

Although measuring the same concepts, the four NHS questions differ slightly from the Aboriginal questions on the 2006 Census. The question wording was modified to reflect current terminology and ensure ongoing accuracy when measuring the Aboriginal population. For example, in 2011 changes were introduced to the terminology in the Aboriginal group question (Question 18). 'North American Indian' was replaced by 'First Nations (North American Indian)' and 'Inuit (Eskimo)' was replaced by 'Inuk (Inuit)' in both the question and the response categories. An instruction was added '**Note: First Nations (North American Indian) includes Status and Non-Status Indians.**' See the earlier Aboriginal concepts section (Section 1.1) for the details about each question.

6.1.2 Legislative changes

Legislative changes such as the amendments to the *Indian Act* of Canada: Bill C-31 in 1985 and Bill C-3 in 2011,[20] may affect concepts such as Aboriginal identity and Registered or Treaty Indian status. Such legislative changes can affect how individuals may have responded to the Aboriginal identity and Registered or Treaty Indian status questions in the 2011 NHS. The net effect of these changes cannot be measured. Users should be careful when interpreting the results.

20. Bill C-3, the *Gender Equity in Indian Registration Act*, in effect since January 31, 2011, ensures that eligible grand-children of women who lost status as a result of marrying non-Indian men will become entitled to registration (Indian status). For more information on Bill C-3, refer to www.aadnc-aandc.gc.ca/eng/1305747570701/1305747904278.

6.1.3 Differences in the list set of incompletely enumerated reserves

In 2011, there were a total of 36 Indian reserves and Indian settlements reported[21] as 'incompletely enumerated' in the NHS. For 18 reserves or settlements, census enumeration was either not permitted or was interrupted before it could be completed and so the NHS was not administered in those areas. For four (4) reserves or settlements, census enumeration was completed however, data collection for the NHS was not permitted or interrupted and in one reserve or settlement, it was determined that there was no resident population contrary to what was erroneously reported in the census. In the case of 13 reserves in Northern Ontario, enumeration was delayed because of natural events (specifically forest fires) and estimates for these communities are not included in geographic areas that include these communities (e.g., provincial and national estimates).

There are no data for incompletely enumerated Indian reserves and settlements on the NHS database. Higher-level geographic areas containing these areas are identified in the NHS products. Although NHS data are not available for incompletely enumerated Indian reserves and settlements, the areas themselves are included as part of the standard geographic hierarchies on the NHS databases.

The extent of the impact of the exclusion of the incompletely enumerated reserves will depend on the geographic area under study. It is much less for higher geographic areas such as Canada, provinces and territories, census metropolitan areas, and census agglomerations. The impact may be more significant for lower geographic areas, such as census subdivisions.

The issue of incompletely enumerated reserves has the most impact on NHS estimates for the First Nations population living on reserve as well as the Registered or Treaty Indian status population living on reserve. Estimates for other populations will also be affected. Other affected populations include: total Aboriginal identity population, First Nations (North American Indian) identity population, total Aboriginal ancestry population, First Nations (North American Indian) ancestry population, Registered or Treaty Indian status population, population who reported membership in a First Nation/Indian band, and population living on Indian reserves and Indian settlements.

Estimates associated with other variables related to First Nations, such as language and band housing, may also be affected by the incomplete enumeration of certain Indian reserves and Indian settlements in the NHS. This issue does not have an impact on the estimates for the Inuit or Métis populations as they are much less likely to live on reserve.

6.1.4 Changes made to the census subdivision types associated with 'on reserve' population

Statistics Canada uses the definition of 'on reserve' provided by Aboriginal Affairs and Northern Development Canada (AANDC). From time to time, there are changes made to the geographies that define Indian reserves and settlements. In 2006, there were 1,174 'on reserve' census subdivisions (CSD) while in 2011, there were 997. The changes are the result of additions, deletions, changes in geographic coding and amalgamations. See the Census subdivision (CSD) definition in the 2011 *Census Dictionary* for the reserves associate with each type of change. Because these are geographic changes, adjustments for these are possible.

Users need to request special tabulations for data that have been adjusted for these changes in order to compare populations on the affected geographies (e.g., comparing on- and off-reserve population).

21. For more information on 2011 NHS incompletely enumerated Indian reserves and settlements, refer to www12.statcan.gc.ca/nhs-enm/2011/ref/aboriginal-autochtones-eng.cfm.

6.1.5 Processing and methodological changes

6.1.5.1 Editing and imputation

The editing and imputation (E & I) of the Aboriginal variables was almost entirely redesigned for 2011, with the primary goal being to streamline the processes and to use, as much as possible, one donor to impute data for a respondent who had provided incomplete or invalid responses on his/her NHS questionnaire. In 2011, the variables of immigration, citizenship, place of birth, ethnic origin/Aboriginal ancestry, population group/visible minority, Aboriginal group, Registered or Treaty Indian status, and First Nation/Indian band membership were processed together, with the interrelations between these variables clearly defined in advance. Donor imputation for missing information within these variables was done with one donor for all variables, as much as possible.

Given the low item non-response and invalid rates, and the correspondingly low imputation rates for Aboriginal identity group, Registered or Treaty Indian status, and First Nation/Indian band membership, the overall effect of E & I on the data quality of these variables is not nearly as significant as other processes, such as the weighting procedures used to deal with non-responding households.

6.1.5.2 Weighting and calibration

The NHS weighting process involves calculating sampling weights, adjusting the weights for the survey's total non-response and calibrating the weights against census totals.

First, an initial sampling weight of about 3 is assigned to each sampled household. However, the sampling fraction varies with the questionnaire delivery mode. For the mail delivery mode, about 3 in 10 households (29%) received a questionnaire. For the enumerator delivery mode, the sampling fraction was 1 in 3 households (33%). Nevertheless, in cases where it was necessary to reach households in remote areas or on Indian reserves, where only the interview response mode was offered, all households were invited to participate in the NHS.

Then the sampling weights are adjusted to reflect the selection of the subsample from the set of households that had not responded to the NHS by mid-July 2011. Next, since a number of households in the subsample were still non-respondents at the end of collection operations, the sampling weight was adjusted for the survey's residual non-response. This was done by transferring the weights of non-respondent households to the nearest-neighbour-respondent households.

Lastly, the weights are calibrated against census totals at the level of geographic calibration areas. Calibration is performed so that the estimates for an NHS calibration area are approximately equal to the census counts for that area, for a set of about 60 characteristics common to the NHS and the census.

Nevertheless, there may be differences between the NHS estimates and the census counts for common characteristics. The smaller the geographic area is, the greater the risk that the NHS estimates will be different from the census counts. This was present with the 2006 Census long form, but it was less common because of the higher response rates and the small variation in these response rates across areas, for both small and large municipalities.

For more information on the methodology of the NHS, refer to the *National Household Survey User Guide*, Catalogue no. 99-001-X2011001. All efforts are made to reduce errors in estimation and the Census of Population plays a major role in ensuring the reliability in the estimates of the NHS.

6.2 Undercoverage of population in participating Indian reserves and settlements

In 2011, it is estimated that the net undercoverage of the population living in participating Indian reserves and settlements is 10,125 persons or 2.8%, compared to 2.2% net undercoverage for the remaining population in the 2011 Census. For detailed information on issues related to coverage, such as overcoverage and undercoverage, see the earlier coverage section (Section 5.3).

The specific undercoverage of the off-reserve Aboriginal population for 2011 is not available.

6.3 Comparability with other sources

6.3.1 Comparability with the Indian Register

The 2011 National Household Survey (NHS) from Statistics Canada, as well as the censuses of population previous to that, and the Indian Register maintained by Aboriginal Affairs and Northern Development Canada (AANDC) are the principal sources of demographic data on the Registered (or Status) Indian population.

The 2011 NHS, like censuses before 2011, are snapshots of the population that self-identify as Registered or Treaty Indians at a moment in time, while the Indian Register is a continuous statutory administrative file based on the registration of individuals who meet specific criteria as defined by the *Indian Act*.

The NHS estimate is affected by undercount of Registered Indians living on incompletely enumerated reserves where enumeration was is not permitted, or was interrupted before it could be completed; undercoverage of Registered Indians living on participating reserves and settlements; and undercoverage of Registered Indians living off reserve. As well, the NHS did not collect data on Registered Indians living in institutions (for example, hospitals, senior citizens' homes, jails, shelters, etc.), nor did the NHS collect data on Registered Indians living outside of the country on NHS collection day.

The Indian Register, on the other hand, forms one cluster of AANDC's entire Indian Registration System (IRS). While considered the authoritative source for the number of Registered Indians as per criteria established in the *Indian Act*, the Indian Register also comes with its own set of limitations. First of all, the main purpose of the Indian Register is to record individual names and a range of non-statutory data in accordance with specific subsections of the *Indian Act*. It was not designed for statistical analysis and therefore data from the application may not fully meet the requirements of some statistical activities (e.g., demographic projections, migration patterns).

There is a reporting lag between the occurrence of a given life event and its being reported and recorded for updating of the Indian Register. This means that events occurring in a given year may not be reflected in the Indian Register during that year. The following examples taken from the *Registered Indian Population by Sex and Residence 2011* report illustrate this concept.

1. Infants entitled to be registered at birth may not be registered by the December 31st reporting date. Parents often do not report the birth until a later year. It is estimated that of all births reported in 2011, approximately 73% had actually occurred prior to 2011.

2. Individuals can remain on the Indian Register for some time after they are deceased. A certificate of death or a confirmation of presumed death is normally required to remove a name from the Indian Register.

3. Residency codes are typically updated when a life event is reported, although some bands update them more frequently. Residency remains a voluntary field in that the information is not mandatory to collect. For additional information on the Indian Register, refer to www.aadnc-aandc.gc.ca/eng/1351001356714/1351001514619.

In addition, Registered Indian status does not guarantee Canadian residency. On the Indian Register there are a number of registrants living in other countries. For example, at the 2011 year-end, there were nearly 19,000 registrants whose province of residence was 'Outside of Canada.'

The different purposes, as well as methodological and conceptual differences between the Indian Register and the NHS, result in estimates that are not directly comparable. The estimate of Registered Indians from the NHS, on May 10, 2011, is 697,510 which is 19.7% lower than the estimate from the Indian Register (IR) as of December 31, 2011 (868,206).

6.3.2 Comparability with Demosim projections

Demosim is a microsimulation model developed and maintained at Statistics Canada, and designed for population projections. Using the microdata file from the Canadian Census of Population (20% sample) as its starting point, Demosim produces dynamic population projections at the level of the provinces, territories, census metropolitan areas and selected smaller geographies, based on a number of characteristics such as age, sex, visible minority group, place of birth, generation status, Aboriginal identity, highest level of educational attainment and labour force participation, among others. It does so by simulating events such as births, deaths, migrations and changes in level of education, according to various population growth scenarios.[22]

Different Demosim scenarios were prepared specifically for the purposes of comparing NHS and projection results for the Aboriginal variables: Aboriginal identity and Registered or Treaty Indian status. For Aboriginal identity variable, two scenarios were prepared: one scenario that does not include ethnic mobility[23] and keeps Aboriginal fertility rates constant; and another scenario that includes ethnic mobility according to the average from 1996 to 2001 and 2001 to 2006, and Aboriginal fertility rates are kept constant. A third scenario does not include ethnic mobility, Aboriginal fertility rates are kept constant and there is no migration on the reserves.

---

22. For more information on the Demosim model, refer to the methodology sections of the following documents: *Projections of the Diversity of the Canadian Population, 2006 to 2031* (www.statcan.gc.ca/pub/91-551-x/2010001/meth-eng.htm) and *Population Projections by Aboriginal Identity in Canada, 2006 to 2031* (www.statcan.gc.ca/pub/91-552-x/2011001/meth-eng.htm). Together, the two sections, complementing each other, provide additional information about Demosim's content, including its starting population, its general functioning and the data sources and methods used to model its components.
23. Ethnic mobility refers to the change in reporting of Aboriginal identity over time, from one census period to the next. In the recent past, ethnic mobility has been shown to be a major component in the growth of the Métis population and to a lesser extent, First Nations population (Guimond, 1999; Guimond, 2003; Lebel, Caron Malenfant and Guimond, 2011).

At the Canada level, the 2011 NHS Aboriginal identity results (1,400,690) are very close to those of the Demosim projection ethnic mobility scenario (1,393,775) – a difference of 6,915 or 0.5%. However, the difference increases to 102,580 (7.9%), when comparing the NHS results with those of the scenario without ethnic and/or cultural mobility. Furthermore, the difference between the two sources is more noticeable for the different Aboriginal groups, especially the Métis and the First Nations (North American Indian).

Ethnic and/or cultural mobility is the transference of identity group that has been observed when comparing one census to the next (more so for non-Aboriginal identity to an Aboriginal identity). Its effect on the estimates of First Nations and Métis are noted more than for the estimates of Inuit. Consequently, overall we observe that the differences in estimates between the two sources are smaller between the NHS and the Demosim scenario with ethnic and/or cultural mobility, than between NHS and the Demosim scenario without ethnic and/or cultural mobility.

7 Conclusions

In summary, the NHS provides an essential part of the information on Aboriginal people in Canada. There are no other sources of information that can provide detailed socioeconomic information for such small domains of interest. The Aboriginal population estimates may change over time due to demographic growth and for a number of other non-demographic reasons. Users of NHS data on Aboriginal people should be aware of changes that affect the interpretation of the data, especially when comparing between past censuses and the 2011 National Household Survey.

For concepts and variables, it was noted that although measuring the same concepts, the four NHS questions differ slightly from the comparable Aboriginal questions on the 2006 Census. The question wording was modified to reflect current terminology and ensure ongoing accuracy when measuring the Aboriginal population.

There were changes noted in a number of survey and statistical processes; among them sampling, collection, processing and estimation. The NHS had a number of differences from the 2006 Census long form, for example in terms of the voluntary nature of the survey , the sampling fraction, collection strategies and response options (e.g., more people used the Internet option) and a lower response rate including different approach to adjust for non-response.

The coverage of the NHS also differs from the past. The universe for the 2011 NHS represents a different population than the one covered by the 2006 Census. The target population for the 2006 Census long form includes usual residents in collective dwellings and persons living abroad whereas the target population for the NHS excludes them. The list of incompletely enumerated Indian reserves and Indian settlement is also different. Finally, unlike previous collection periods, the NHS collection of the 13 Indian reserves or Indian settlements in Northern Ontario is a unique circumstance where survey data has been collected where the census enumeration was not done.

Other changes that affected the results were outside the survey data collection process. Legislative changes, for example Bill C-31 in 1985 and Bill C-3 in 2011, which affect concepts such as Aboriginal identity and Registered or Treaty Indian status and changes made to the definition of reserves, will have an impact on the interpretation of the information collected.